EDGEWOOD

CHEMICAL BIOLOGICAL CENTER

U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND

ECBC-TR-857

DISCRIMINATION OF PATHOGENIC VS. NONPATHOGENIC *FRANCISELLA TULARENSIS* AND *BURKHOLDERIA PSEUDOMALLEI* USING PROTEOMICS MASS SPECTROMETRY

Mary Margaret Wade
Charles H. Wick
Alan W. Zulich

RESEARCH AND TECHNOLOGY DIRECTORATE

Rabih E. Jabbour

SCIENCE APPLICATIONS
INTERNATIONAL CORPORATION
Gunpowder, MD 21010-0068

Samir V. Deshpande

SCIENCE AND TECHNOLOGY CORPORATION
Edgewood, MD 21040-2734

March 2011

ABERDEEN PROVING GROUND, MD 21010-5424

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
XX-03-2011	Final	Oct 2009 - Sep 2010

4. TITLE AND SUBTITLE
Discrimination of Pathogenic vs. Nonpathogenic *Francisella tularensis* and *Burkholderia pseudomallei* Using Proteomics Mass Spectrometry

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Wade, Mary Margaret; Wick, Charles H.; Zulich, Alan W. (ECBC); Jabbour, Rabih E. (SAIC); and Deshpande, Samir V. (STC)

5d. PROJECT NUMBER
None

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DIR, ECBC, ATTN: RDCB-DRB-D, APG, MD 21010-5424
SAIC, P.O. Box 68, Gunpowder, MD 21010-0068
STC, 500 Edgewood Road, Suite 205, Edgewood, MD 21040-2734

8. PERFORMING ORGANIZATION REPORT NUMBER
ECBC-TR-857

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DIR, ECBC, ATTN: RDCB-D, In-house Laboratory Independent Research Program, APG, MD 21010-5424

10. SPONSOR/MONITOR'S ACRONYM(S)
ILIR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT-LIMIT 200 WORDS

The recent use of mass spectrometry (MS) analysis to characterize and identify biological agents using a proteomic approach has been demonstrated. The present study sought to determine whether proteomics MS could be used to distinguish between pathogenic and nonpathogenic strains of the same organism based on outer membrane proteins (OMPs). OMPs are often associated with virulence in gram-negative pathogens and could prove to be excellent model biomarkers for strain differentiation among bacteria. OMPs were isolated from pathogenic and nonpathogenic strains of *Francisella tularensis* and *Burkholderia pseudomallei* using ultracentrifugation and a sarkosyl extraction method.These methods are followed by protein digestion and analysis using liquid chromatography-MS/MS. The experimental MS/MS spectral data of bacterial peptides were searched using the SEQUEST algorithm against a constructed proteome database of microorganisms to determine the identity and number of unique peptides. Results showed that pathogenic strains could be distinguished from nonpathogenic strains for both organisms based on differences in OMP profiles. Virulent *F. tularensis* and *B. pseudomallei* expressed multiple OMPs that were not present in avirulent samples. In conclusion, proteomics MS proved useful in discriminating between pathogenic and non-pathogenic strains, and OMPs unique to pathogenic strains could prove to be useful biomarkers for strain differentiation.

15. SUBJECT TERMS
Francisella tularensis | *Burkholderia pseudomallei* | Outer membrane proteins | Proteomics | Mass spectrometry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Sandra J. Johnson
U	U	U	UL	27	19b. TELEPHONE NUMBER *(include area code)* (410) 436-2914

Blank

PREFACE

The work described in this report was started in October 2009 and completed in September 2010.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for public release. Registered users should request additional copies from the Defense Technical Information Center. Unregistered users should direct such requests to the National Technical Information Service.

Acknowledgment

This work was funded by the 2010 U.S. Army Edgewood Chemical Biological Center's In-house Laboratory Independent Research program.

Blank

CONTENTS

FIGURES

TABLES

DISCRIMINATION OF PATHOGENIC VS. NONPATHOGENIC *FRANCISELLA TULARENSIS* AND *BURKHOLDERIA PSEUDOMALLEI* USING PROTEOMICS MASS SPECTROMETRY

1. INTRODUCTION

The recent use of mass spectrometry (MS) analysis has proven useful in characterizing and identifying biological agents using a proteomic approach (1). Therefore, the present study sought to determine whether proteomics MS could be used to distinguish between pathogenic and nonpathogenic strains of the same organism. More specifically, discrimination between pathogenic and nonpathogenic organisms based on their outer membrane protein (OMP) composition, as determined by MS, was investigated.

OMPs of gram-negative bacteria act as active mediators between the cell and its environment and are often associated with virulence in gram-negative pathogens (2-6). Avirulent strains often lack one or more of the plasmids or genes encoding proteins needed for virulence. These differences in OMP expression between virulent and avirulent strains of gram-negative bacteria could potentially be exploited to distinguish among strains. Therefore, OMPs could prove to be excellent model biomarkers for strain differentiation among bacteria.

The objective of the present study was to establish the sequence-based identity of OMPs isolated from pathogenic and nonpathogenic strains of *Francisella tularensis* and *Burkholderia pseudomallei*. *F. tularensis* is classified as a Category A pathogen, is an important potential biowarfare agent (BWA), and is the causative agent of tularemia. *B. pseudomallei* is classified as a Category B pathogen and is considered a potential BWA. *B. pseudomallei* is the causative agent of melioidosis and infects humans and animals. Although not endemic to the United States like *F. tularensis*, *B. pseudomallei* is a risk to soldiers serving in areas such as Southeast Asia and Australia where the disease is known to occur.

This one year basic research study aimed to: (1) isolate OMPs using ultracentrifugation and differential extractions, (2) determine sequence and post-translational modifications to amino acid residues composing membrane proteins using emerging high-throughput mass spectral proteomic systems, and (3) use bioinformatics modeling tools to establish strain differentiation methods based on the proteome differences among the *F. tularensis* and *B. pseudomallei* strains.

2. MATERIALS AND METHODS

2.1 Materials and Reagents.

Ammonium bicarbonate, dithiotheritol, urea, acetonitrile-HPLC grade, and formic acid were purchased from Burdick and Jackson (St. Louis, MO). Sequencing grade-modified trypsin was purchased from Promega (Madison, WI).

2.2 Bacterial Strains and Culture Conditions.

Pathogenic strains used in the present study were *F. tularensis* Schu4 and *B. pseudomallei* 1026B. Nonpathogenic strains used were *F. tularensis* LVS and *B. thailandensis* E264. Working cultures were prepared by streaking cells from cryopreserved stocks onto chocolate agar (CA) for *F. tularensis* strains and nutrient agar (NA) for *Burkholderia* strains, followed by incubation for approximately 18 h at 37°C. After incubation, all working culture plates were stored at 4°C. Cells from working cultures were used to inoculate broth cultures for each strain, which consisted of 100 ml of brain heart infusion (BHI) supplemented with cysteine (1% final concentration) for *F. tularensis* and 100 ml of nutrient broth (NB) for *Burkholderia*. All cultures were incubated for approximately 18 h at 37°C with rotary aeration at 180 rpm. After incubation, broth cultures were pelleted by centrifugation (2,300 RCF at 4°C for 10 min), washed, and resuspended in 10 ml HEPES buffer followed by heating at 95°C for 1 h to lyse cells. After heating, a portion of each sample was plated onto appropriate agar as described above and incubated for 5 days at appropriate temperatures previously listed to ensure no growth prior to removing samples from either the BSL-2 or BSL-3 laboratory for further processing. Total cellular protein samples (whole cell lysates) were complete after heating for 1 h and were transferred to Point Detection Branch for analysis after no growth on plates was confirmed. Prior to being transferred to Point Detection Branch for analysis, OMP samples were processed for OMP isolation as discussed in Section 2.3.

2.3 OMP Isolation.

After lysis by heating at 95°C for 1 h, cell debris was pelleted by centrifugation at 2,300 RCF at 4°C for 10 min. The supernatant was then centrifuged at 100,000 x *g* for 1 h to pellet proteins. The pellet was resuspended in 1 ml of HEPES buffer. One milliliter of a 2% sarkosyl solution (N-Lauroylsarcosine sodium salt solution) was added, and the sample was incubated at room temperature for 30 min. Next, samples were centrifuged at 100,000 x *g* for 1 h; the pellet containing OMPs was resuspended in 1 ml of HEPES buffer, and then transferred to the Point Detection Branch for further processing and analysis as discussed in Section 2.4.

2.4 Processing of Whole Cell Lysates and OMP Samples.

All protein samples were ultra-sonicated (20 s pulse on, 5 s pulse off, and 25% amplitude for 5 min duration; Branson 450 digital sonifier; Branson, Danbury, CT) and a small portion of lysates was reserved for 1-D gel analysis. The lysates were centrifuged at 14,100 x *g* for 30 min to remove any debris. The supernatant was then added to a Microcon YM-3 filter unit (Millipore; Cat #: 42404) and centrifuged at 14,100 x *g* for 30 min. The effluent was discarded. The filter membrane was washed with 100 mM ammonium bicarbonate (ABC) and centrifuged for 15-20 min at 14,100 x *g*. Proteins were denatured by adding 8 M urea and 3 µg/µl Dithiotheritol (DTT) to the filter and incubating overnight at 37°C on an orbital shaker set to 60 rpm. Twenty microliters of 100% acetonitrile (ACN) was added to the tubes and allowed to incubate at room temperature for 5 min. The tubes were then centrifuged at 14,100 x *g* for 30-40 min and washed three times using 150 µl of 100 mM ABC solution. On the last wash, ABC was allowed to sit on the membrane for 20 min while shaking, followed by centrifugation at 14,100 x *g* for 30-40 min. The micron filter unit was then transferred to a new receptor tube,

and proteins were digested with 5 µl trypsin in 240 µl of ABC solution + 5 µl ACN. Proteins were digested overnight at 37°C on an orbital shaker set to 55 rpm. Sixty microliters of 5% ACN/0.5% formic acid (FA) was added to each filter to quench the trypsin digestion followed by 2 min of vortexing for sample mixing. The tubes were centrifuged for 20-30 min at 14,100 x *g*. An additional 60 µl 5% ACN/0.5% FA mixture was added to filter and centrifuged. The effluent was then analyzed using the LC-MS/MS technique.

2.5 Protein Database and Database Search Engine.

A protein database was constructed in a FASTA format using the annotated bacterial proteome sequences derived from fully sequenced chromosomes of 881 bacteria, including their sequenced plasmids (as of April 2009). A PERL program (http://www.activestate.com/Products/ActivePerl; accessed April 2009) was written to download these sequences automatically from the National Institutes of Health National Center for Biotechnology (NCBI) site (http://www.ncbi.nlm.nih.gov; accessed April 2009). Each database protein sequence was supplemented with information about a source organism and a genomic position of the respective ORF embedded into a header line. The database of bacterial proteomes was constructed by translating putative protein-coding genes and consists of tens of millions of amino acid sequences of potential tryptic peptides obtained by the *in silico* digestion of all proteins (assuming up to two missed cleavages).

The experimental MS/MS spectral data of bacterial peptides were searched using SEQUEST algorithm against a constructed proteome database of microorganisms. The SEQUEST thresholds for searching the product ion mass spectra of peptides were Xcorr, deltaCn, Sp, RSp, and deltaMpep. These parameters provided a uniform matching score of all candidate peptides. The generated out files of these candidate peptides were then validated using a peptide prophet algorithm. Peptide sequences with a probability score of 95% and higher were retained in the dataset and used to generate a binary matrix of sequence-to-bacterium assignments. The binary matrix assignment was populated by matching the peptides with corresponding proteins in the database and assigning a score of 1. A score of 0 was assigned for a nonmatch. The column in the binary matrix represents the proteome of a given bacterium, and each row represents a tryptic peptide sequence from the LC-MS/MS analysis. Microorganisms were matched with the bacterium/bacteria based on the number of unique peptides that remained after further filtering of degenerate peptides from the binary matrix. Verification of the classification and identification of candidate microorganisms was performed through hierarchical clustering analysis and taxonomic classification.

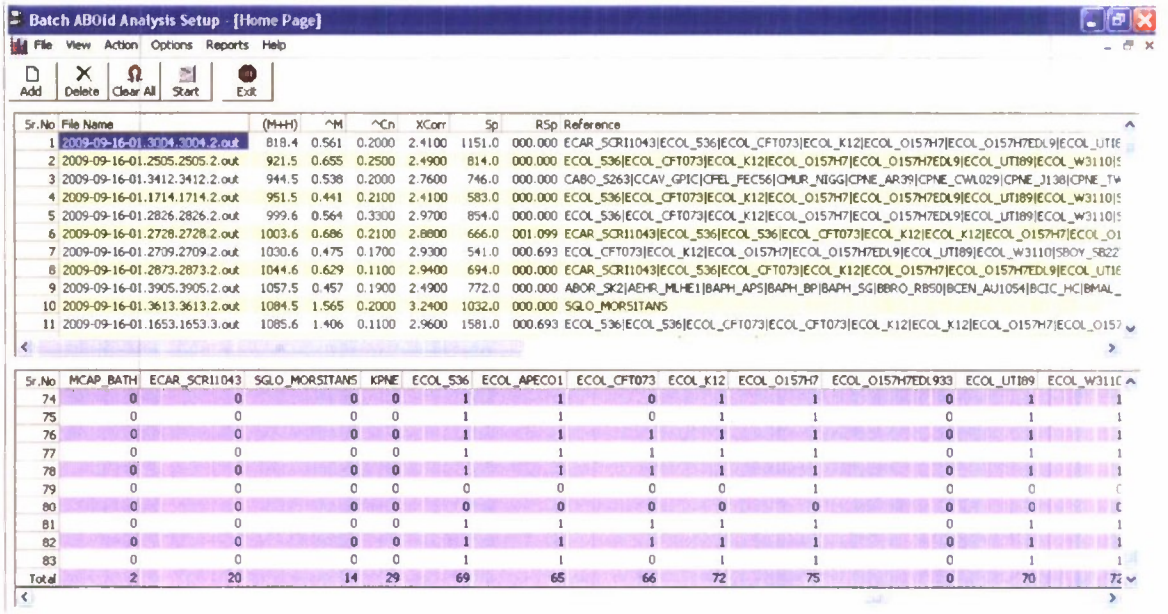
The in-house developed software called "BACid" transformed results of searching MS/MS spectra of peptide ions against a custom protein database, which was downloaded from NCBI with commercial software SEQUEST into a taxonomically meaningful and easy to interpret output. It calculated probabilities that peptide sequence assignment to a MS/MS spectrum was correct and used accepted spectrum–to-sequence matches to generate a sequence-to-bacterium (STB) binary matrix of assignments. Validated peptide sequences, differentially present or absent in various strains (STB matrices), were visualized as assignment bitmaps and analyzed by a BACid module that used phylogenetic relationships among bacterial species as a part of the decision tree process. The bacterial classification and identification

algorithm used assignments of organisms to taxonomic groups (phylogenetic classification) based on an organized scheme that begins at the phylum level and follows through classes, orders, families, and genus down to the strain level. BACid was developed in-house using PERL, MATLAB, and Microsoft Visual Basic.

3. RESULTS AND DISCUSSION

The current project characterized and identified pathogenic and nonpathogenic strains of the same organism based on proteins present in whole cell lysates (global) versus OMP preparations (specific). All results are shown and discussed below.

Figure 1 serves as an example to illustrate the typical output generated for the LC-ESI MS/MS analyses of bacterial proteins digest using bioinformatics tools to process the peptide sequence information for the bacterial differentiation and classification. The top window lists the identified unique proteins and their corresponding bacterium match. The lower window represents the binary matrix of the sequence-to-bacterium search matching. The total row, lower window, represents the total number of unique proteins identified for a given bacterium. Figure 2 also serves as an example and shows the histogram generated by plotting the number of unique proteins versus the bacterium matching in the database. The Y-axis represents the percentage of unique peptides matched with 95% confidence level for all the bacteria on the x-axis. In this example case, the identified bacterium, which is at the strain level, is *Y. pestis*. The horizontal redline is the threshold cutoff under which common degenerate peptides among various bacteria within the constructed proteome database are shown. These degenerate peptides are removed from the total number of unique peptides of the identified species.



Sr.No	File Name	(M+H)	^M	^Cn	XCorr	Sp	RSp	Reference
1	2009-09-16-01.3004.3004.2.out	818.4	0.561	0.2000	2.4100	1151.0	000.000	ECAR_SCR11043|ECOL_536|ECOL_CFT073|ECOL_K12|ECOL_O157H7|ECOL_O157H7EDL9|ECOL_UTI
2	2009-09-16-01.2505.2505.2.out	921.5	0.655	0.2500	2.4900	814.0	000.000	ECOL_536|ECOL_CFT073|ECOL_K12|ECOL_O157H7|ECOL_O157H7EDL9|ECOL_UTI89|ECOL_W3110|
3	2009-09-16-01.3412.3412.2.out	944.5	0.538	0.2000	2.7600	746.0	000.000	CABO_S263|CCAV_GPIC|CFEL_FEC56|CMUR_NIGG|CPNE_AR39|CPNE_CWL029|CPNE_J138|CPNE_TW
4	2009-09-16-01.1714.1714.2.out	951.5	0.441	0.2100	2.4100	583.0	000.000	ECOL_536|ECOL_CFT073|ECOL_K12|ECOL_O157H7|ECOL_O157H7EDL9|ECOL_UTI89|ECOL_W3110|
5	2009-09-16-01.2826.2826.2.out	999.6	0.564	0.3300	2.9700	854.0	000.000	ECOL_536|ECOL_CFT073|ECOL_K12|ECOL_O157H7|ECOL_O157H7EDL9|ECOL_UTI89|ECOL_W3110|
6	2009-09-16-01.2728.2728.2.out	1003.6	0.686	0.2100	2.8800	666.0	001.099	ECAR_SCR11043|ECOL_536|ECOL_536|ECOL_CFT073|ECOL_K12|ECOL_K12|ECOL_O157H7|ECOL_O1
7	2009-09-16-01.2709.2709.2.out	1030.6	0.475	0.1700	2.9300	541.0	000.693	ECOL_CFT073|ECOL_K12|ECOL_O157H7|ECOL_O157H7EDL9|ECOL_UTI89|ECOL_W3110|SBOY_SB22
8	2009-09-16-01.2873.2873.2.out	1044.6	0.629	0.1100	2.9400	694.0	000.000	ECAR_SCR11043|ECOL_536|ECOL_CFT073|ECOL_K12|ECOL_O157H7|ECOL_O157H7EDL9|ECOL_UTI
9	2009-09-16-01.3905.3905.2.out	1057.5	0.457	0.1900	2.4900	772.0	000.000	ABOR_SK2|AEHR_MLHE1|BAPH_APS|BAPH_BP|BAPH_SG|BBRO_RB50|BCEN_AU1054|BCIC_HC|BMAL_
10	2009-09-16-01.3613.3613.2.out	1084.5	1.565	0.2000	3.2400	1032.0	000.000	SGLO_MORSITANS
11	2009-09-16-01.1653.1653.3.out	1085.6	1.406	0.1100	2.9600	1581.0	000.693	ECOL_536|ECOL_536|ECOL_CFT073|ECOL_CFT073|ECOL_K12|ECOL_K12|ECOL_O157H7|ECOL_O157

Sr.No	MCAP_BATH	ECAR_SCR11043	SGLO_MORSITANS	KPNE	ECOL_536	ECOL_APEC01	ECOL_CFT073	ECOL_K12	ECOL_O157H7	ECOL_O157H7EDL933	ECOL_UTI89	ECOL_W311
74	0	0	0	0	1	1	0	1	1	0	1	1
75	0	0	0	0	1	1	0	1	1	0	1	1
76	0	0	0	0	1	1	1	1	1	0	1	1
77	0	0	0	0	1	1	1	1	1	0	1	1
78	0	0	0	0	1	1	1	1	1	0	1	1
79	0	0	0	0	0	0	0	0	1	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0
81	0	0	0	0	1	1	1	1	1	0	1	1
82	0	0	0	0	1	1	1	1	1	0	1	1
83	0	0	0	0	1	1	0	1	1	0	1	1
Total	2	20	14	29	69	65	66	72	75	0	70	72

Figure 1. MS-based proteomic approach output. The upper section represents the matching algorithm results of the identified tryptic peptides resulting from the LC-MS/MS analysis. The lower section represents the binary matrix of STB scoring. Presence of a unique peptide corresponding with a protein in the given proteome of a bacterium is scored 1, and a nonmatch is scored 0.

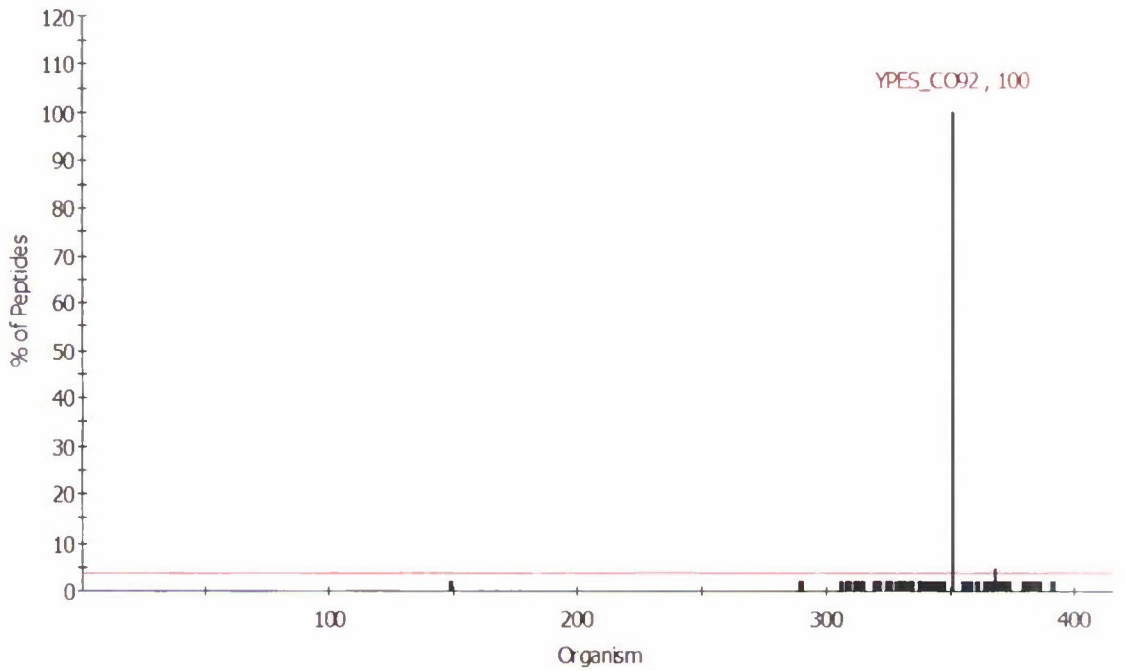


Figure 2. Histogram representing the output of the binary matrix of the unique peptides identified for a given bacterium at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

3.1 Differentiation of Pathogenic vs. Nonpathogenic *F. tularensis* Strains Using Whole Cell Lysates and OMPs.

Pathogenic and nonpathogenic *F. tularensis* Schu S4 and LVS, respectively, were analyzed by proteomic MS for identification using OMPs and whole cell lysates. Figures 3 through 6 show the histogram for the STB binary matrix, with the number of unique peptides on the y-axis and bacterium proteome on the x-axis. As seen in the figures, correct identification of each strain was achieved.

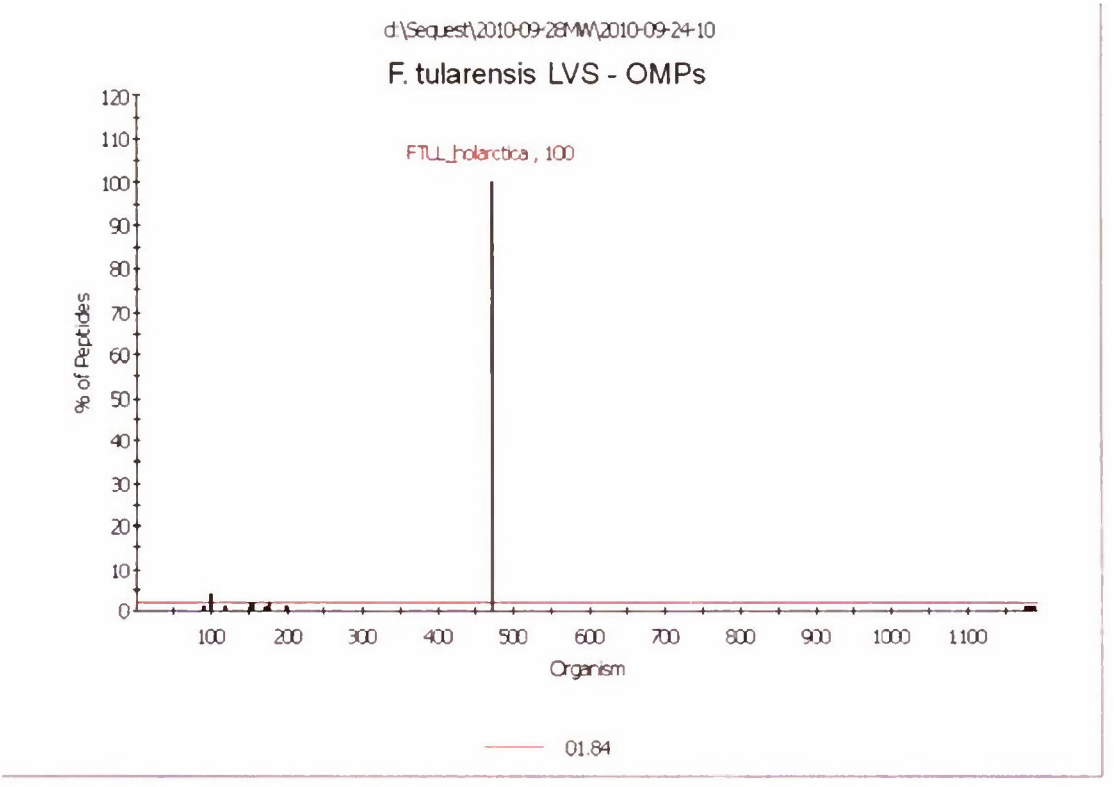


Figure 3. Bacterial differentiation of *F. tularensis* LVS using OMPs. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

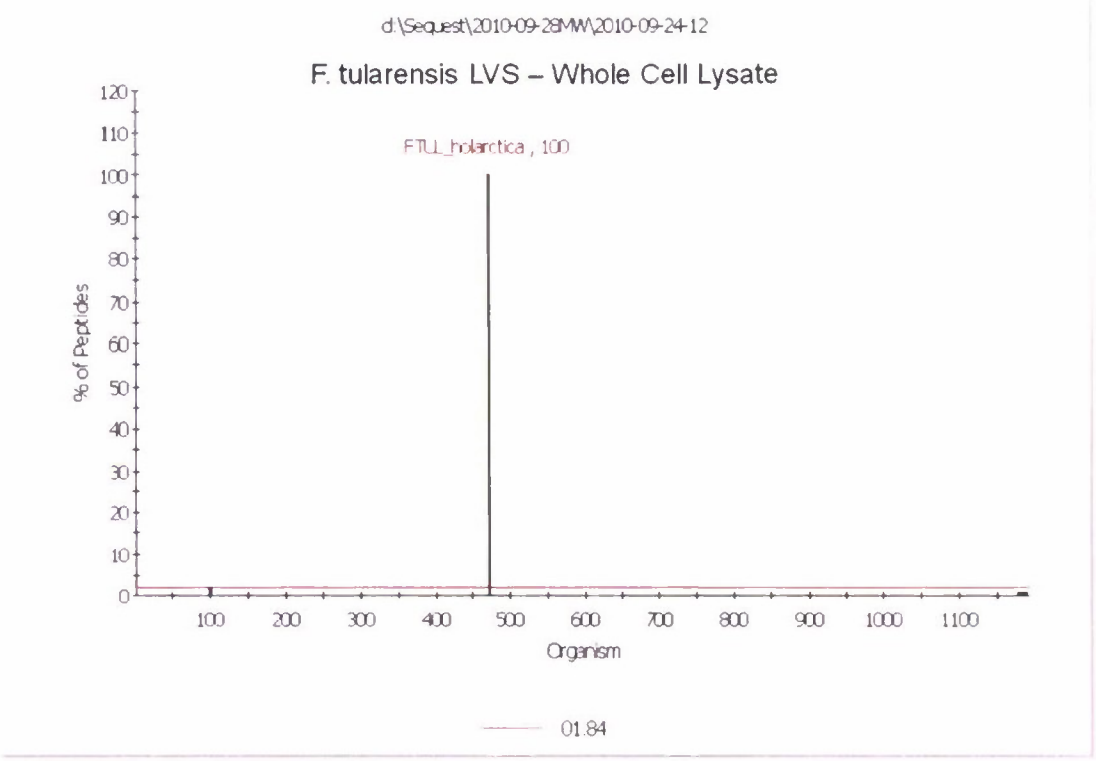


Figure 4. Bacterial differentiation of *F. tularensis* LVS using Whole Cell Lysates. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

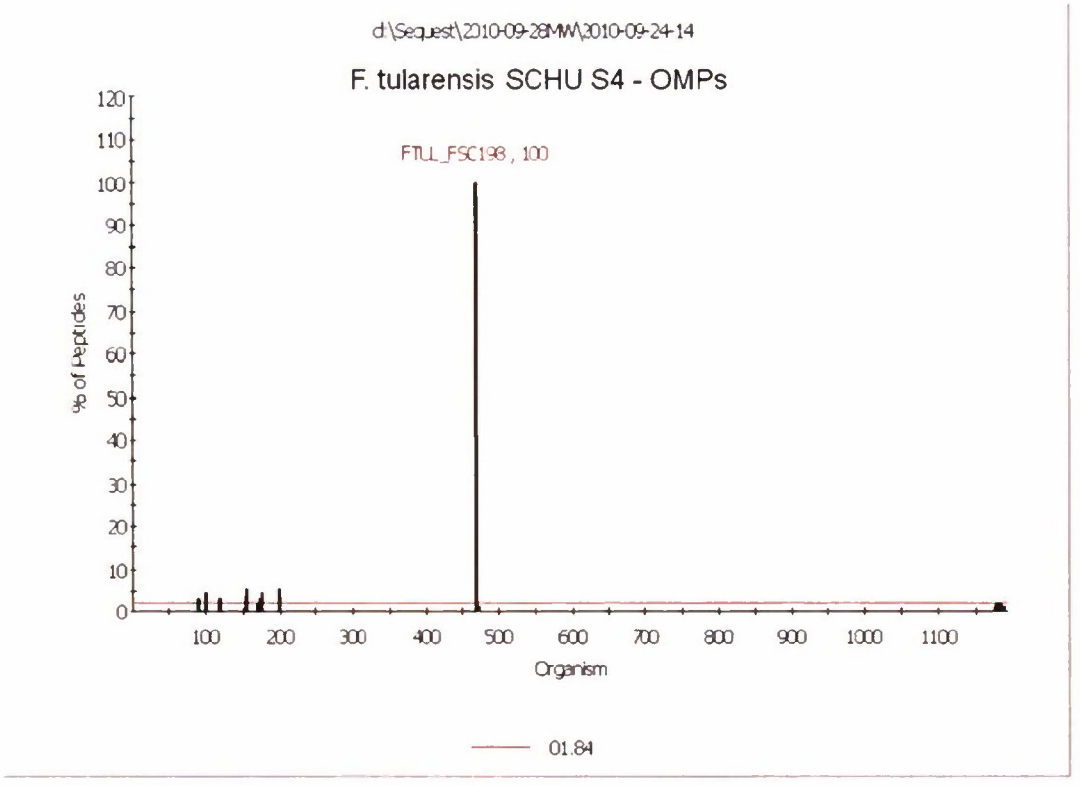


Figure 5. Bacterial differentiation of *F. tularensis* Schu S4 using OMPs. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

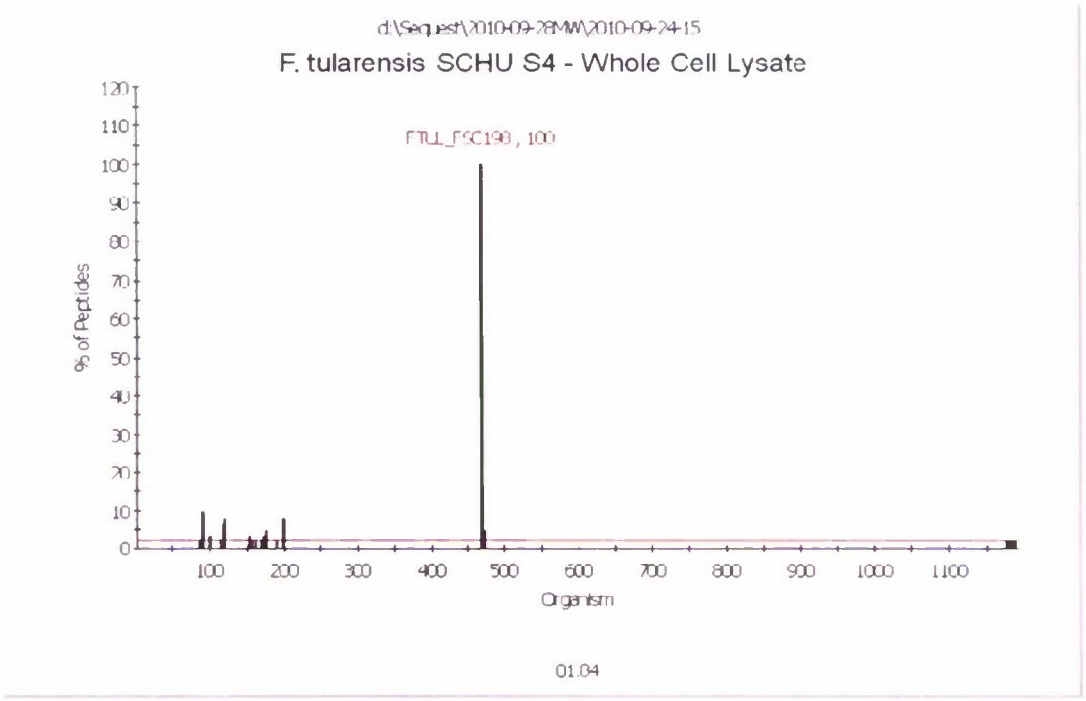


Figure 6. Bacterial differentiation of *F. tularensis* Schu S4 using Whole Cell Lysates. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

OMPs of pathogenic and nonpathogenic *F. tularensis* Schu S4 and LVS, respectively, were prepared and analyzed by proteomic MS as described above. Results showed correct identification at the strain level for both samples analyzed. The near neighbor analysis, using the Euclidean distance linkage approach, for these lysed bacterial samples showed that the identified unique set of proteins had the closest match with the used *F. tularensis* strains. Therefore, correct identification to the strain level was achieved for both bacterial OMPs (Figures 7 and 8).

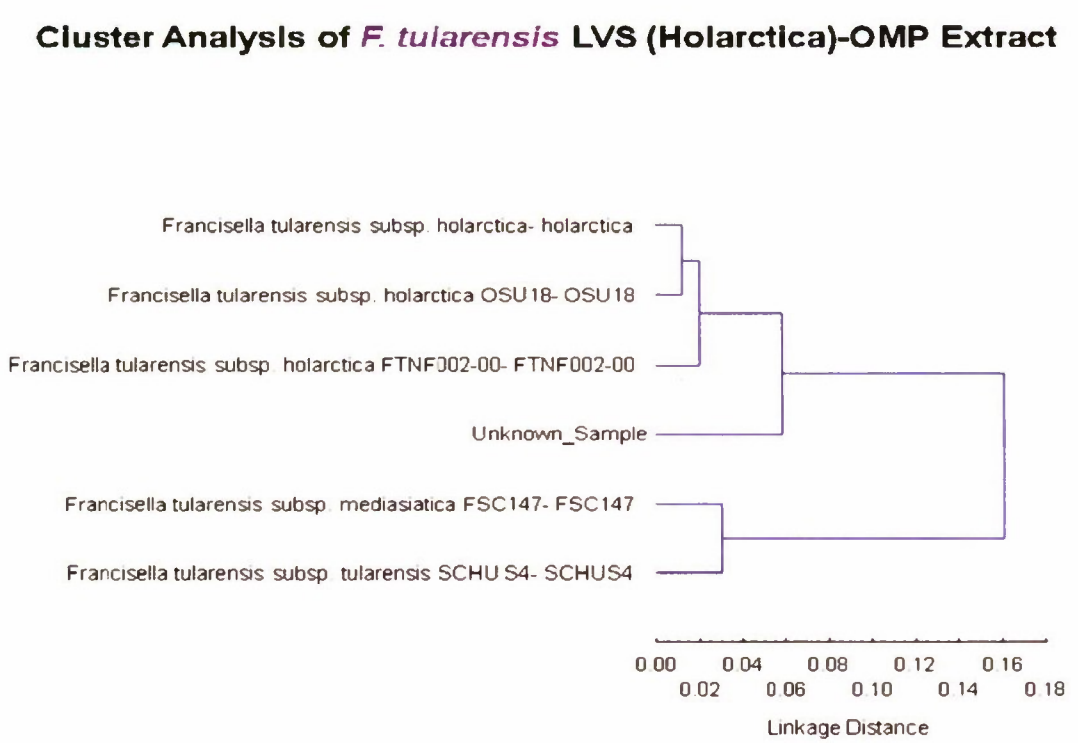


Figure 7. Near-neighbor classification of nonpathogenic *F. tularensis* LVS using OMPs.

Cluster Analysis of *F. tularensis* *SCHUS4*-OMP Extract

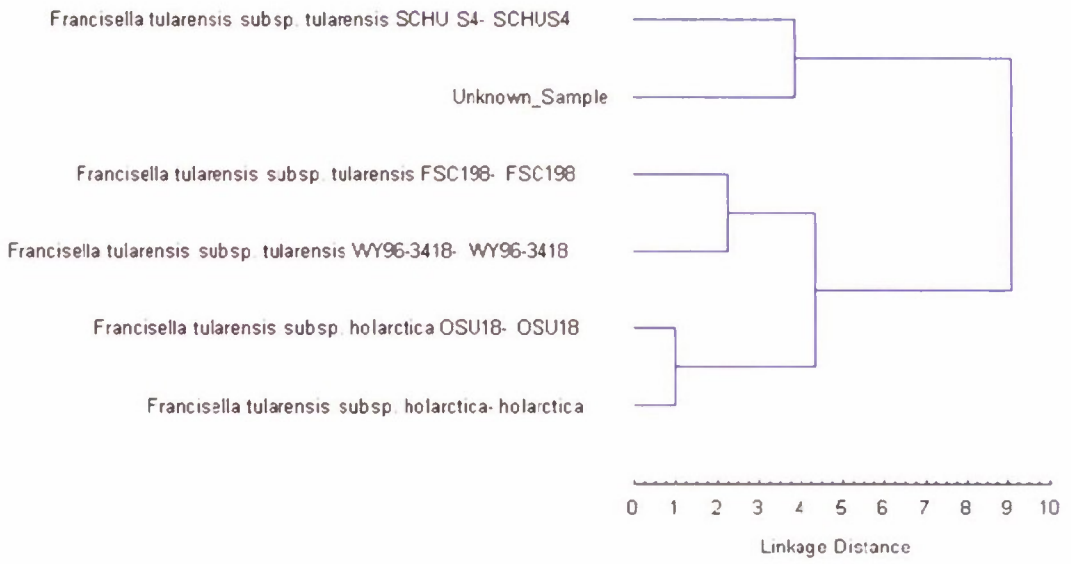


Figure 8. Near-neighbor classification of pathogenic *F. tularensis* Schu S4 using OMPs.

Table 1 shows the comparison of the strain unique proteins for *F. tularensis* Schu S4 and LVS identified from the different cellular extracts. The number of unique proteins identified differed between the pathogenic and nonpathogenic *F. tularensis* strains, with the pathogenic strain having a relatively lower number of unique proteins than that of the nonpathogenic *F. tularensis*. In addition, comparing whole cell lysate to OMP extract showed a variation in the number of strain unique protein biomarkers in OMPs versus whole cell lysates. The number of strain unique proteins from OMPs versus whole cell lysate was slightly higher for pathogenic versus nonpathogenic strains, and nonpathogenic strain LVS had a slightly higher number of unique proteins from OMPs.

Table 1. Unique Proteins for *F. tularensis* Schu S4 and *F. tularensis* LVS from Whole Cell Lysates vs. OMPs.

F. tularensis Schu S4 Unique Proteins	
Whole Cell Lysates	OMPs
Type IV pili fiber building block protein hypothetical protein FTH_1232 OMP hypothetical protein FTT_1334c chaperonin GroEL	Type IV pili fiber building block protein molecular chaperone DnaK 3-oxoacyl-[acyl-carrier-protein] synthase II lipoprotein chaperonin GroEL heat shock protein 90 OMP
F. tularensis LVS Unique Proteins	
Whole Cell Lysates	OMPs
chaperonin GroEL hypothetical protein FTL_0617 AhpC/TSA family protein heat shock protein DnaK hypothetical protein FTL_0572 hypothetical protein FTL_1110 hypothetical protein FTL_1494 macrophage infectivity potentiator, fragment OMP co-chaperonin GroES	outer membrane associated protein chaperonin GroEL hypothetical protein FTL_0617 succinyl-CoA synthetase subunit beta intracellular growth locus, subunit C sigma-54 modulation protein OMP AhpC/TSA family protein heat shock protein DnaK 3-oxoacyl-[acyl-carrier-protein] synthase II lipoprotein chaperonin GroEL co-chaperonin GroES DNA-directed RNA polymerase subunit beta intracellular growth locus, subunit C

3.2 Differentiation of Pathogenic vs. Nonpathogenic *B. pseudomallei* Strains Using Whole Cell Lysates and OMPs.

Pathogenic *B. pseudomallei* 1026B and nonpathogenic *B. thailandensis* E264 were analyzed by proteomic MS for identification using OMPs and whole cell lysates. Figures 9 through 12 show the histogram for the STB binary matrix, with the number of unique peptides on the y-axis and bacterium proteome on the x-axis. As seen in the figures, correct identification of each strain was achieved.

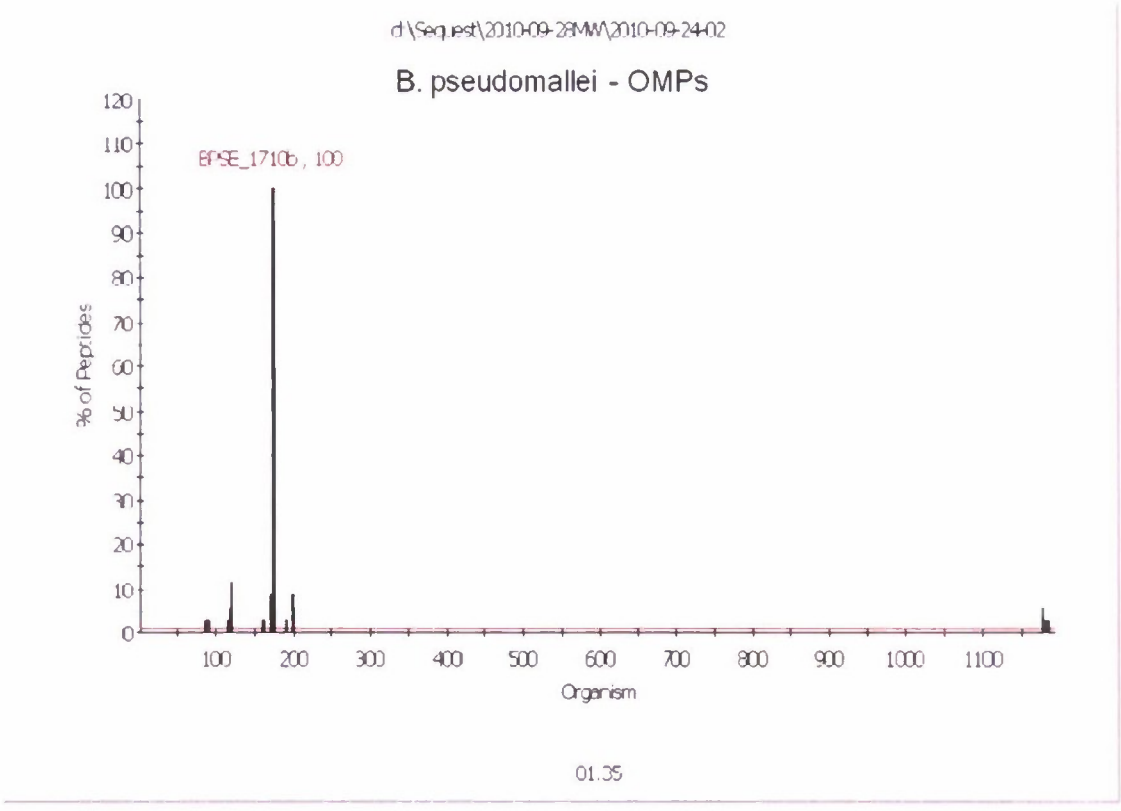


Figure 9. Bacterial differentiation of *B. pseudomallei* 1026B using OMPs. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

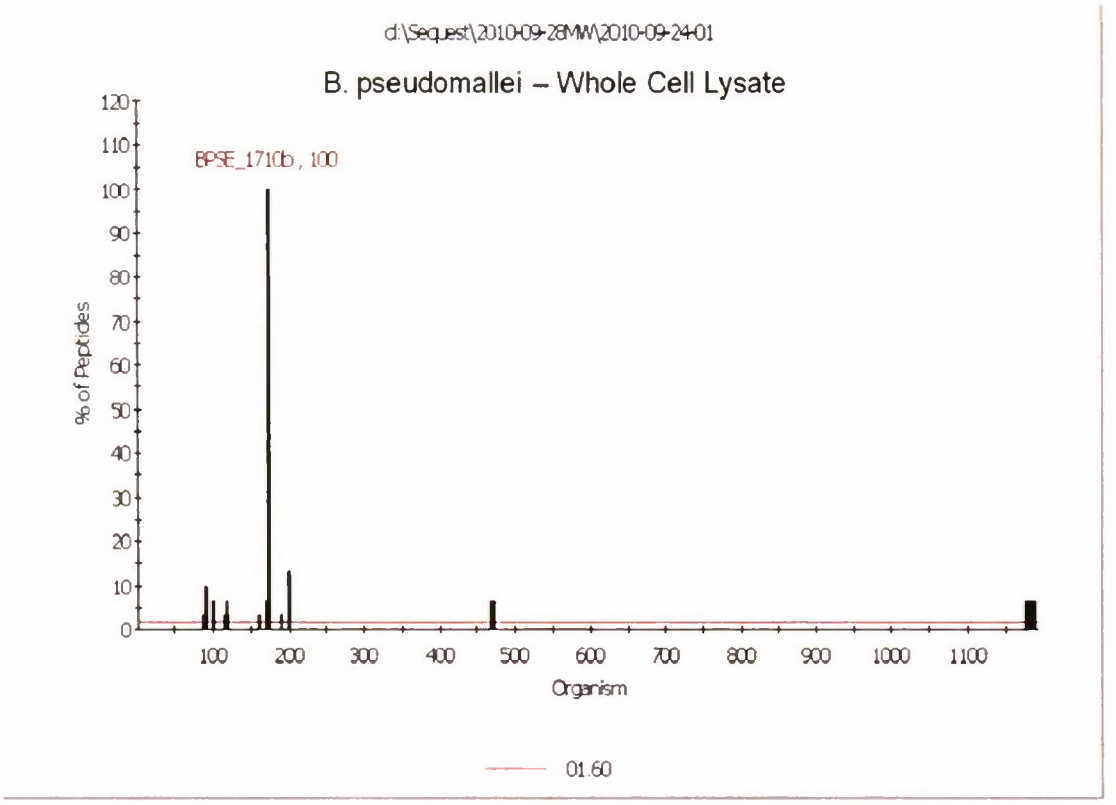


Figure 10. Bacterial differentiation of *B. pseudomallei* 1026B using Whole Cell Lysates. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

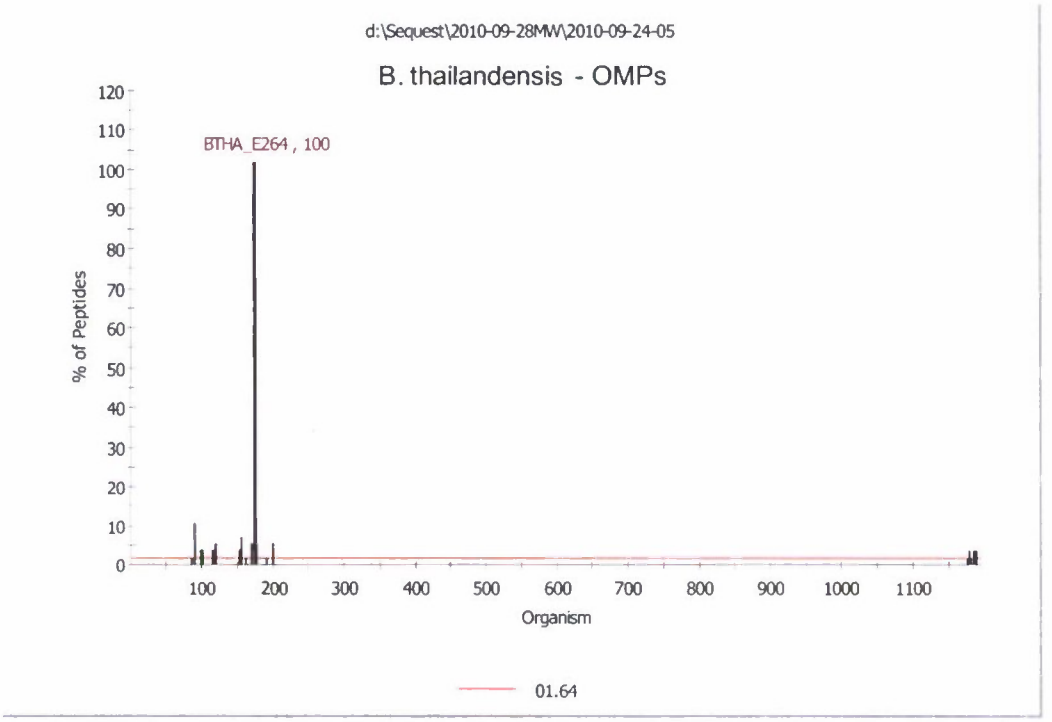


Figure 11. Bacterial differentiation of *B. thailandensis* E264 using OMPs. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

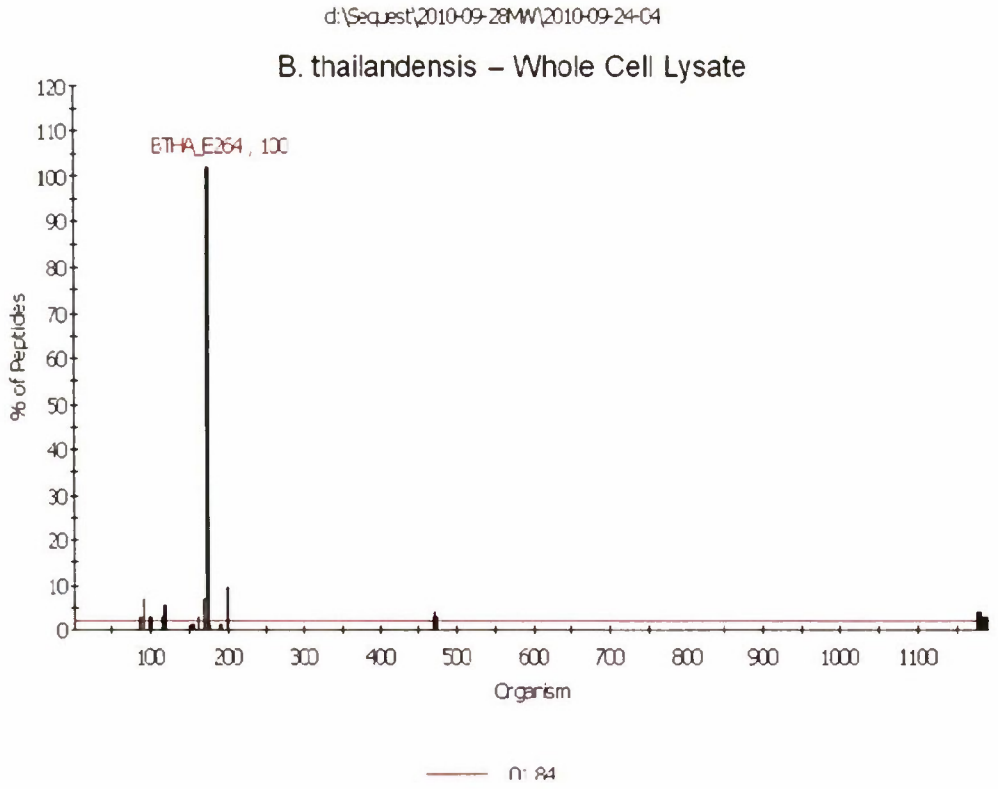


Figure 12. Bacterial differentiation of *B. thailandensis* E264 using Whole Cell Lysates. The X-axis represents the bacterium proteome, and the Y-axis represents the number of unique peptides at 95% confidence level. The horizontal line is the threshold under which peptides identified are considered statistically nonsignificant.

OMPs of pathogenic and nonpathogenic *B. pseudomallei* 1026B and *B. thailandensis* E264, respectively, were prepared and analyzed by proteomic MS as described above. Results show correct identification at the strain level for both samples analyzed. The near neighbor analysis, using Euclidean distance linkage approach, for these lysed bacterial samples showed that the identified unique set of proteins had the closest match with the used *B. pseudomallei* and *B. thailandensis* strains. Therefore, correct identification to the strain level was achieved for both bacterial OMPs (Figures 13 and 14).

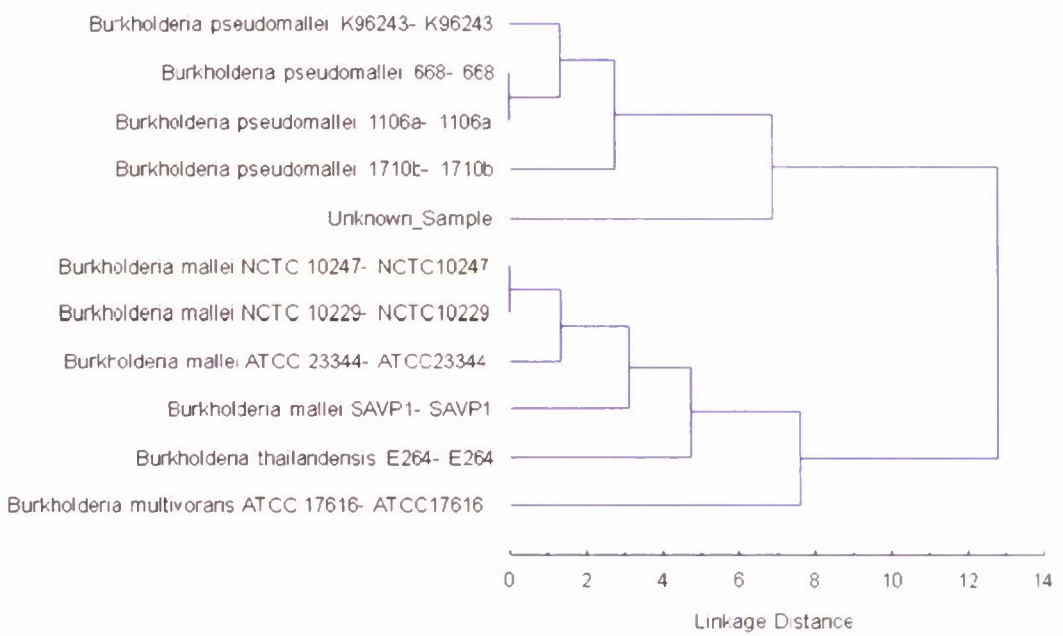


Figure 13. Near-neighbor classification of pathogenic *B. pseudomallei* 1026B using OMPs.

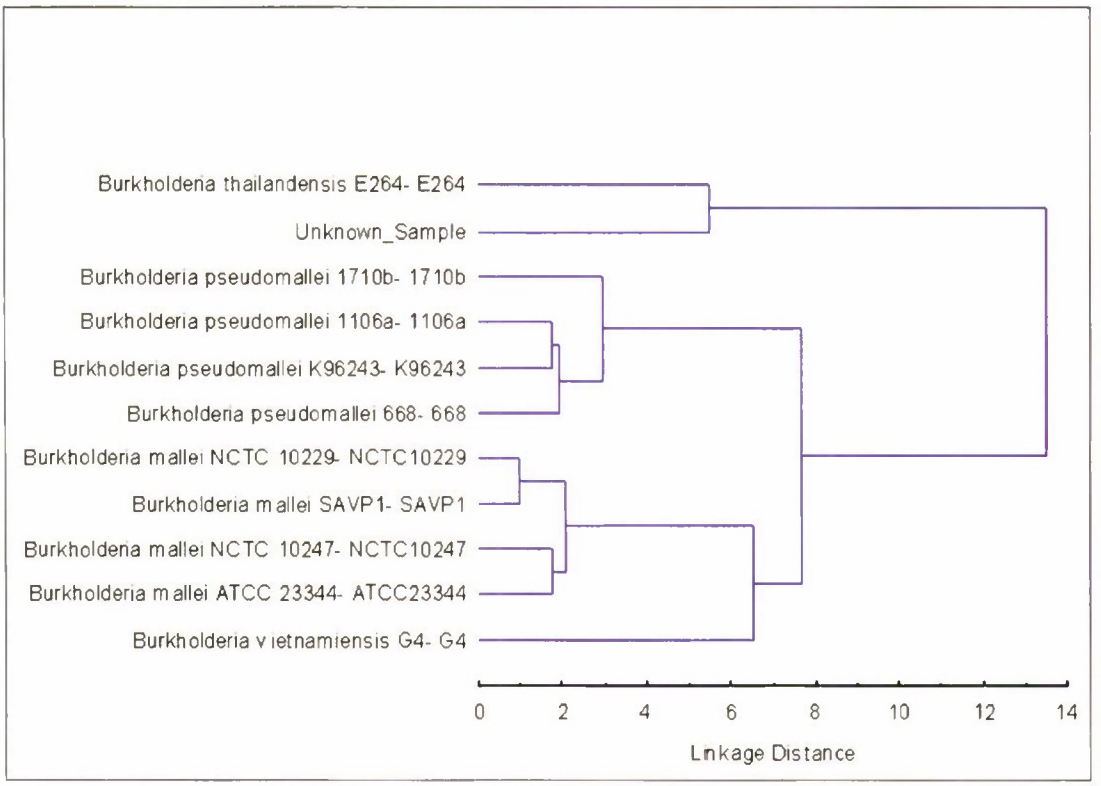


Figure 14. Near-neighbor classification of nonpathogenic *B. thailandensis* E264 using OMPs.

Table 2 shows the comparison of the strain unique proteins for *B. pseudomallei* and *B. thailandensis* identified from the different cellular extracts. The number of unique proteins identified differed between the pathogenic and nonpathogenic strains, with the pathogenic strain having a relatively lower number of unique proteins than that of nonpathogenic *B. thailandensis*. In addition, comparing whole cell lysate versus OMP extract showed a variation in the number of strain unique protein biomarkers in OMPs versus whole cell lysates. The number of strain unique proteins was higher for nonpathogenic *B. thailandensis* from OMPs extracts than whole cell lysate.

Table 2. Unique Proteins for *B. pseudomallei* and *B. thailandensis* from Whole Cell Lysates vs. OMPs.

B. pseudomallei 1026 B Unique Proteins	
Whole Cell Lysates	OMPs
pyochelin synthctase hypothetical protein BPSL2599 putative cytochrome C precursor-related protein	ABC transport system, ATP-binding protein hypothetical protein BPSS2201 glutamate/aspartate periplasmic binding protein precursor hypothetical protein BURPS1710b_2354 hypothetical protein BURPS1710b_2705
B. thailandensis E264 Unique Proteins	
Whole Cell Lysates	OMPs
putative lipoprotein glucan 1,4- a-glucosidase glycosy hydrolase family protein putative lipoprotein glutamate/aspartate ABC transporter, periplasmic glutamate/aspartate-binding protein putative lipoprotein RND efflux system outer membrane lipoprotein putative lipoprotein sensor histidine kinase, putative putative lipoprotein chaperonin GroEL	stress response protein serine protease, MucD putative lipoprotein hypothetical protein BTH_I0089 glutamate/aspartate ABC transporter, periplasmic glutamate/aspartate-binding protein hypothetical protein BTH_II2182 putative lipoprotein gamma-glutamyltransferase putative lipoprotein transport-associated domain-containing protein chaperonin GroEL amino acid ABC transporter, periplasmic amino acid-binding protein, putative dihydrolipoamide succinyltransferase hypothetical protein BTH_II2183 hypothetical protein BTH_I0090 hypothetical protein BTH_I2265 hypothetical protein BTH_I1688 NADPH-dependent FMN reductase domain-containing protein OmpA family protein amino acid ABC transporter, periplasmic amino acid-binding protein translation initiation factor IF-2 ribonuclease E alcohol dehydrogenase, zinc-containing

4. CONCLUSIONS

This project revealed the advantage of using outer membrane proteins (OMPs) as unique biomarkers for bacterial differentiation of pathogenic versus nonpathogenic strains. The differentiation capability enhanced the confidence level of the discrimination process using OMPs as biomarkers. OMPs provide a unique source of cellular variability. Thus, they introduce biodiversity among cellular proteins for very similar bacterial strains, thereby providing distinct and unique protein biomarkers. Although OMPs offered enhanced differentiation capability, whole cell lysates did provide discrimination and are an appropriate option for the differentiation of gram-positive bacterial strains. Overall, an extension of this project, to include a wider investigation of other relevant pathogenic bacteria, could provide us with a more global outlook of the importance of OMPs with regard to pathogenicity. In addition, an extension could increase our confidence in identifying organisms at the strain level using protein biomarkers for bacterial classification and diagnostic purposes.

LITERATURE CITED

1. Dworzanski, J.P.; Deshpande, S.V.; Chen, R.; Jabbour, R.E.; Snyder, A.P.; Wick, C.H.; Lim, L. 2006. Mass spectrometry-based proteomics combined with bioinformatic tools for bacterial classification. *J. Proteome Res.* 5: 76.

2. Nikaido, H. 1996. Outer membrane, p. 29-47. *In* F.C. Neidhardt et al. (cd.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, DC.

3. Haake, D.A. 2000. Spirochaetal lipoproteins and pathogenesis. *Microbiology* 146:1491-1504.

4. Hancock, R.E.; Siehnel, R.; Martin, N. 1990. Outer membrane proteins of Pseudomonas. *Mol. Microbiol.* 4:1069-1075.

5. Schuch, R.; Maurelli, A.T. 1999. The mxi-Spa type III secretory pathway of *Shigella flexneri* requires an outer membrane lipoprotein, MxiM, for invasion translocation. *Infect. Immun.* 67:1982-1991.

6. Zhang, H.; Niesel, D.W.; Peterson, J.W.; Klimpel, G.R. 1998. Lipoprotein release by bacteria: potential factor in bacterial pathogenesis. *Infect. Immun.* 66:5196-5201.